Information Theoretical Measure for Career Determination

ASHNIL MANDALIYA[1], MANOJ SAHNI[2], RAJKUMAR VERMA[3]

[1]School of Liberal Studies,
Pandit Deendayal Petroleum University, Gandhinagar
Gujarat, INDIA

[2]Department of Mathematics,
School of Technology
Pandit Deendayal Petroleum University, Gandhinagar
Gujarat, INDIA.

[3]Department of Management Control and Information Systems
University of Chile,
Av. Diagonal Paraguay 257,
Santiago-8330015, CHILE.

{ashnilman, manojsahani117, rkver83}@gmail.com

Abstract: - A lot of career opportunities are available for high school students. But choosing a right career becomes a most difficult task. The objective of this paper is to help students to choose suitable career after completion of High School. It is based on the student's perception of their own and their Teacher's marks using generalized intuitionistic fuzzy divergence measure. The intuitionistic fuzzy divergence measure is used in table to predict careers suitable for students using the aggregated intuitionistic fuzzy values. The aggregated marks of students and teachers perception are converted in the form of intuitionistic fuzzy values and are shown in table. The final table shows the fuzzy divergence values and the minimum value in the table will be the preferable choice of the students to choose career depending on the marks given by students and teachers perception in six subjects – Maths, Physics, Chemistry, Biology, Computer and English.

Key-Words: - Fuzzy sets, Intuitionistic Fuzzy Sets, Aggregator operator, Divergence Measure, Career Determination.

1 Introduction

The theory of sets is based on the concept that the elements either are included or excluded from the set, i.e., based upon whether they follow the rule or they do not. But in day-to-day life that is not the case; elements may belong to a set to a certain degree. There may also be vagueness in the definition of the rule. Zadeh, in 1965[1], extended the concept of crisp set theory to that of fuzzy set theory wherein elements have a degree of membership or inclusiveness and non-membership, i.e. non- inclusiveness, whose sum is equal to 1. This degree of inclusiveness and non-inclusiveness also known as the membership value ranges from 0 to 1. Fuzzy sets with membership value 1 or 0 are crisp sets. Thus, fuzzy set is a generalization of classical sets. There is a wide range of applications of Fuzzy set theory such as multi-criteria decision making, image processing, pattern recognition, traffic monitoring system, etc.

This concept was extended by Atanassov in 1986 [2] to Intuitionistic fuzzy sets (IFS) to overcome the difficulties face by Zadeh's fuzzy set theory. It takes into account hesitancy which is missing in Fuzzy set theory. He demonstrated with an example where it is shown that in some situations IFS can be applied and fuzzy sets concept cannot be applied [3]. In A-IFS, there are membership and non-membership values assigned to each element. Hesitancy value is given by subtracting the sum of membership and non-membership values from 1. As the name suggests, membership value represents the degree of inclusiveness, non-membership value represents the degree of exclusiveness and hesitancy value represents the uncertainty of inclusiveness or exclusiveness. As crisp sets are a part of fuzzy sets, fuzzy sets are a part of intuitionistic fuzzy sets. A-IFS is applied wherever FS can be applied and, in some fields, it improves the results due to the introduction of hesitancy value such as multi criteria decision making, medical decision support system, psychological analysis, etc.[4-8]. The concept of distances in intuitionistic fuzzy sets was defined by Szmidt et al. [9-10]. The new definitions of distances introduced in early 20th century are compared with already available distances in fuzzy sets. In fuzzy sets, distances are defined using membership and non-membership values, where as in A-IFS also includes the hesitancy values. A distance measure between two Intuitionistic fuzzy sets was defined by Hatzimichailidis et al. [11] in 2012. It is a generalization of the previous distances proposed and the author has shown application of it in pattern recognition. The results obtained are accurate from the already known distances and has shown high degree of confidence.

Aggregation operators were developed to conjugate two sets as the use of Max and Min operators which caused the loss of data and hence filled the gap in real data and aggregated data. It was first introduced in 1985 by D. Dubois and H. Prade [12] where they defined arithmetic mean and geometric mean for fuzzy sets. Zu [13] in 2007 have extended the concept of aggregation operator and defined Intuitionistic Fuzzy Hybrid Averaging (IFHA) Operator. Yager [14] extended the concept of weighted aggregation operators to ordered weighted operators (OWA), where the membership values of the elements of the fuzzy set is ordered in descending order. Since then many more aggregation operators have been developed. These operators are widely used in the field of data mining, multi criteria decision making, sensor fusion etc. Verma et al. [15] has proposed a new divergence measure which provides flexibility in multi-criteria decision making and its applications are shown in this paper.

In this paper, IFS concept is used to draw the conclusion of the career based on the students and teachers perception of the subjects chosen by the students. A questionnaire response from the candidate and teacher's is used to convert the data into intuitionistic fuzzy values. The aggregator operator and generalized intuitionistic fuzzy divergence measure [15] is used to analyze the result obtained.

2 Preliminaries

In this section, we will define some basic definitions related to fuzzy sets and Intuitionistic fuzzy sets. We will list the distances and aggregation operators for both fuzzy sets and Intuitionistic Fuzzy Sets. Further the generalized intuitionistic fuzzy divergence measure [15] will be used in this research paper for calculating divergency between two intuitionistic fuzzy set matrices.

Definition 1 (Fuzzy Sets) [1]: Let us consider a non-empty set X. A fuzzy set A defined on the elements of the set X having the membership value $\mu_A(x)$, defined as $A = \{< x, \mu_A(x) >: x \in X\}$

Definition 2 (Intuitionistic Fuzzy Sets) [2]: Let us consider a non-empty set X. An intuitionistic fuzzy set A defined on the elements of the set X having the membership value $\mu_A(x)$ and non-membership value $\nu_A(x)$, defined as $A = \{\langle x, \mu_A(\mathrm{x}), \nu_A(x) \ \rangle: x \in X\}$, where $\mu_A(\mathrm{x}) + \ \nu_A(x)$ lies in the interval [0,1].

Furthermore, we have $\pi_A(x) = 1 - (\mu_A(\mathrm{x}) + \nu_A(x))$ called the hesitancy margin of x in A.

Definition 3 (Distances for FS) [9]: The most widely used distance measures for Fuzzy Sets A, B in $X = \{x_1, x_2, \dots, x_n\}$ are

- The Hamming distance, $d(A, B)$

$$d(A,B) = \sum_{i=1}^{n} |\mu_A(x_i) - \mu_B(x_i)|$$

- The normalized Hamming distance, $l(A, B)$

$$l(A,B) = \frac{1}{n}\sum_{i=1}^{n} |\mu_A(x_i) - \mu_B(x_i)|$$

- The Euclidian Distance,

$$e(A,B) = \sqrt{\sum_{i=1}^{n} (\mu_A(x_i) - \mu_B(x_i))^2}$$

- The normalized Euclidian distance, $q(A, B)$

$$q(A,B) = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (\mu_A(x_i) - \mu_B(x_i))^2}$$

Szmidt and Kacprzyk [9] extended the concept of distances defined in fuzzy sets to Intuitionistic Fuzzy Sets, in which the hesitancy factor is also considered which is missing in the distances defined for fuzzy sets.

Definition 4 (Distances for IFS) [9]: Szmidt and Kacprzyk defined the widely used distances for two Intuitionistic Fuzzy Sets, A and B, for all elements x of a non-empty set X as per following:

- Hamming distance

$$d_{IFS}(A,B) = \frac{1}{2}\sum_{i=1}^{n}(|\mu_A(x) - \mu_B(x)| + |\nu_A(x) - \nu_B(x)| + |\pi_A(x) - \pi_B(x)|)$$

- Normalized Hamming distance

$$q_{IFS}(A,B) = \frac{1}{2n}\sum_{i=1}^{n}(|\mu_A(x) - \mu_B(x)| + |\nu_A(x) - \nu_B(x)| + |\pi_A(x) - \pi_B(x)|)$$

- Euclidian distance

$$e_{IFS}(A,B) = \sqrt{\frac{1}{2}\sum_{i=1}^{n}\left(\mu_A(x) - \mu_B(x)\right)^2 + \left(\nu_A(x) - \nu_B(x)\right)^2 + (\pi_A(x) - \pi_B(x))^2}$$

- Normalized Euclidian distance

$$l_{IFS}(A,B) = \sqrt{\frac{1}{2n}\sum_{i=1}^{n}\left(\mu_A(x) - \mu_B(x)\right)^2 + \left(\nu_A(x) - \nu_B(x)\right)^2 + (\pi_A(x) - \pi_B(x))^2}$$

Definition 5 (Intuitionistic Fuzzy Aggregation Operator [IFWA]) [13]: Let $a_j = \left[t_{a_j}, 1 - \right.$ *faj, (j=1, 2,...,n)*, be a collection of intuitionistic fuzzy values; then their aggregated value by using the IFWA is defined as $IFWA_w(a_1, a_2, \dots, a_n) = \left[1 - \prod_{j=1}^{n}\left(1 - t_{a_j}\right)^{w_j}, 1 - \prod_{j=1}^{n}\left(f_{a_j}\right)^{w_j}\right]$, where $w = (w_1, w_2, \dots, w_n)^T$ is the weight vector of $a_j, (j = 1,2,\dots,n)$, with $w_j \in [0,1]$ and $\sum_{j=1}^{n} w_j = 1$.

Definition 6 (Generalized Intuitionistic Fuzzy Divergence) [15]: Let A and B are two intuitionistic fuzzy sets defined on $X = \{x_1, x_2, \dots, x_n\}$ having the membership values $\mu_A(x_i)$ and $\mu_B(x_i), i = 1,2,\dots,n$, and non-membership values $\nu_A(x_i)$ and $\nu_B(x_i), i = 1,2,\dots,n$ respectively. The measure of generalized intuitionistic fuzzy divergence denoted by $D_\lambda(A|B)$ and defined between two IFS A and B as

$$D_\lambda(A|B) = \frac{1}{n}\sum_{i=1}^{n}\begin{bmatrix} \mu_A(x_i)log\frac{\mu_A(x_i)}{\lambda\mu_A(x_i)+(1-\lambda)\mu_B(x_i)} \\ +\nu_A(x_i)log\frac{\nu_A(x_i)}{\lambda\nu_A(x_i)+(1-\lambda)\nu_B(x_i)} \\ +\pi_A(x_i)log\frac{\pi_A(x_i)}{\lambda\pi_A(x_i)+(1-\lambda)\pi_B(x_i)} \end{bmatrix}$$

where $0 \leq \lambda \leq 1$.

Definition 7 (Symmetric Generalized Intuitionistic Fuzzy Divergence) [15]: The symmetric generalized Intuitionistic Fuzzy divergence measure between two IFS A and B is defined as

$$D_\lambda(A;B) = D_\lambda(A|B) + D_\lambda(B|A).$$

In our paper, we have used Generalized Intuitionistic Fuzzy divergence measure as it's generalization of all the previous known measures and its simplicity and applicability in multi-criteria decision making that was shown in the paper.

3 Career Determination after High School Using Fuzzy Divergence Measure

Many research scientists have used the concept of distances in fuzzy sets and Intuitionistic fuzzy sets in various fields such as in research questionnaire, career determination, selection of best student, student evaluation, etc. In this paper, we have formed a Questionnaire and given to six different High School Students and their Teachers of different subjects. The response of the Questionnaire is converted into Intuitionistic fuzzy values in a normalized way and is depicted in tables 1 and 2. The membership and non-membership values are depicted in Table 1 and 2, whereas hesitancy index is calculated as $1 - (\mu + \nu)$. Table 3 is formed from by taking the aggregation values of both Students and Teachers perception. The hypothetical table 4 is created from the IFS values required in each subject versus career. Using the definitions 6 and 7, the generalized Intuitionistic Fuzzy divergence measure between students and careers are depicted in Table 5. The least divergency measure in columns of Table 5 for careers against the students will give choice to the students to choose proper career.

Table 1. Student Perception about High School Result

		Students											
Subjects		S1		S2		S3		S4		S5		S6	
		μ1	ν1	μ2	ν2	μ3	ν3	μ4	ν4	μ5	ν5	μ6	ν6
	Maths	.8	.1	.9	.05	.6	.3	.55	.3	.7	.2	.75	.15
	Physics	.8	.15	.8	.1	.6	.3	.5	.2	.7	.2	.8	.1
	Chemistry	.6	.2	.7	.2	.7	.1	.7	.2	.6	.2	.6	.3
	Biology	.5	.3	.6	.3	.8	.1	.7	.1	.5	.3	.5	.3
	Computer	.8	.1	.7	.2	.6	.2	.5	.2	.9	.05	.85	.1
	English	.7	.2	.6	.2	.7	.2	.8	.1	.7	.2	.7	.2

Table 1 is depicting the student's perception about his / her High School result. Here S1, S2, S3, S4, S5 and S6 are used for six different students. The membership and non-membership values are shown for all students. The hesitancy is calculated using $\pi_i(x) = 1 - (\mu_i(x) + \nu_i(x))$. Similarly Table 2 is drawn taking the perception of the teacher's for students of various subjects. The response is converted into IFS values in a normalized way.

Table 2. Teacher's Perception about the Students

		Students											
Subjects		S1		S2		S3		S4		S5			S6
		μ1	ν1	μ2	ν2	μ3	ν3	μ4	ν4	μ5	ν5	μ6	ν6
	Maths	.75	.15	.8	.15	.55	.25	.5	.2	.65	.3	.8	.1
	Physics	.8	.1	.75	.15	.7	.2	.6	.2	.65	.2	.85	.1
	Chemistry	.6	.2	.7	.2	.8	.15	.7	.2	.6	.1	.7	.2
	Biology	.55	.25	.6	.3	.8	.1	.75	.2	.6	.2	.6	.2
	Computer	.8	.1	.75	.2	.7	.2	.6	.2	.75	.2	.85	.1
	English	.7	.15	.65	.25	.7	.25	.7	.2	.6	.2	.7	.2

Table 3 values are calculated by taking the aggregation of students and Teacher's perception IFS values from Table 1 and 2. Here equal weights are taken for both Table 1 and 2, that is $w_i = 0.5$. Intuitionistic Fuzzy aggregation operator is used to aggregate two tables, i.e. 1 and 2, of concern.

Table 4 is constructed of IFS values in such a manner where more membership value is given depending on the choice of careers. The career in Engineering requires good background of Mathematics, Physics and Computers. Similarly to chose a career in Medical, one should good in Biology, Physics and Chemistry. Similarly, for other careers fuzzy values are given corresponding to each subjects of concern. Table 5 is constructed using the generalized intuitionistic fuzzy divergence measure [15]. It is seen that students S1 and S6 have less value of divergence in Engineering, S2 in B.Sc. Mathematics, S3 in Medical, S4 in B.Sc. Bio, S5 in BCA. Thus, it is concluded from Table 5 that for students S1 and S6 it is better to choose Engineering, S2 to choose B.Sc. Mathematics, S3 to choose Medical, S4 to choose B.Sc. Biology, S5 to choose BCA. So, finally it is concluded from the table that the lesser divergence value will be a preferable choice of the students in choosing career.

Table 3. Aggregation (IFWA) of Table 1 and Table 2

		Students											
Subjects		S1		S2		S3		S4		S5		S6	
		μ_1	ν_1	μ_2	ν_2	μ_3	ν_3	μ_4	ν_4	μ_5	ν_5	μ_6	ν_6
	Maths	.78	.13	.86	.10	.58	.28	.53	.25	.68	.25	.78	.13
	Physics	.80	.13	.78	.13	.65	.25	.55	.20	.68	.20	.83	.10
	Chemistry	.60	.20	.70	.20	.76	.13	.70	.20	.60	.15	.65	.25
	Biology	.53	.28	.60	.30	.8	.10	.73	.15	.55	.25	.55	.25
	Computer	.80	.10	.73	.20	.65	.30	.55	.20	.84	.13	.85	.10
	English	.70	.18	.63	.23	.70	.23	.76	.15	.65	.20	.70	.20

Table 4. IFS values required in each subject versus career

		Subjects											
Career		Maths		Physics		Chemistry		Biology		Computer		English	
		μ_1	ν_1	μ_2	ν_2	μ_3	ν_3	μ_4	ν_4	μ_5	ν_5	μ_6	ν_6
	Engineering	.8	.1	.8	.1	.6	.2	.5	.3	.8	.1	.7	.2
	Medical	.6	.2	.7	.2	.7	.2	.7	.2	.6	.2	.6	.2
	B.Sc. Biology	.5	.3	.6	.2	.7	.2	.7	.2	.5	.3	.7	.2
	B.Sc. Maths	.8	.1	.8	.1	.7	.2	.5	.3	.8	.1	.5	.3
	BCA	.7	.2	.6	.2	.5	.3	.5	.3	.8	.1	.5	.3

Table 5. Careers against Students using Intuitionistic Fuzzy Divergence Measure

		Students					
Career		S1	S2	S3	S4	S5	S6
	Engineering	0.00790	0.03765	0.11746	0.13498	0.04414	0.01735
	Medical	0.22447	0.07087	0.03664	0.03784	0.09085	0.08724
	BSc. Biology	0.39908	0.12438	0.05969	0.02095	0.13003	0.16150
	BSc. Mathematics	0.06868	0.03229	0.13134	0.06725	0.06725	0.02791
	BCA	0.06567	0.07717	0.12954	0.12359	0.04296	0.07140

4 Conclusion

This study provides a method based on IFS theory for students in choosing a subject which may be useful in their future career. In this paper, a generalized intuitionistic fuzzy divergence measure is used to find the proper choice of the career based on the students' and teacher's perception depicted in the form of IFS values in Table 1 and 2. This technique of career determination can be further extended to various generalizations of IFS such as Picture Fuzzy Sets, Pythagorean Fuzzy Sets, Ortho-Pair Fuzzy Sets etc. to improve the accuracy and precision of the proper choice that is determined. Further improvements can be made to the techniques by introducing questionnaires, a subject based and a student-based, to further pinpoint the optimum option while also delving deeper into the intricacies of each subject and career field. Analogous techniques can be developed for different fields including but not limited to Pattern Recognition, Risk Evaluation, Project Evaluation.

References:

[1] Zadeh, L.A., Fuzzy Sets, *Inform. and Control*, Vol. 8, 1965, pp. 338 -353.

[2] Atanassov, K.T., Intuitionistic Fuzzy sets, *Fuzzy Sets Syst.*, Vol. 20, 1986, pp. 87 – 96.

[3] Atanassov, K.T., Intuitionistic Fuzzy Sets Past, Present and Future, *Proceedings of the 3rd Conference of the European Society for Fuzzy Logic and Technology*, Zittau, Germany, September, 2003, pp. 10 – 12.

[4] Ejegwa, P.A., Akubo, A.J., Joshua, O.M., Intuitionistic Fuzzy Set and Its Application in Career Determination via Normalized Euclidean Distance Method, *Eur. Sci. J.*, Vol. 10, No. 15, 2014, pp. 529 – 536.

[5] Yager, R.R., Some Aspects of Intuitionistic Fuzzy Sets, *Fuzzy Optim. Decis. Making*, Vol. 8, 2009, pp. 67 – 90.

[6] Xu Z., Yager, R.R., Dynamic Intuitionistic Fuzzy Multi-attribute Decision Making, *Int. J. Approx. Reason*., Vol. 48, 2008, pp. 246 – 262.

[7] Szmidt, E., Kacprzyk, J., Intuitionistic Fuzzy Sets in Some Medical Applications, Sofia 22-23 Sept. 2001, NIFS 7, Vol. 4, 2001, pp. 58 – 64.

[8] Kozae, A.M., Elshenawy A., Omran, M., Intuitionistic Fuzzy Set and Its Application in Selecting Specialization: A Case Study for Engineering Students, *Int. J. Math. Anal. Appl.*, Vol. 2, No. 6, 2015, pp. 74 – 78.

[9] Szmidt E., Kacprzyk, J., Distances between Intuitionistic Fuzzy Sets, *Fuzzy Sets and Systems*, Vol. 114, No. 3, 2000, pp. 505 – 518.

[10] Szmidt E., Kacprzyk, J., Distances between Intuitionistic Fuzzy sets: A straight forward approaches may not work," 3rd International IEEE Conference Intelligent Systems, 2006.

[11] Hatzimichailidis, A.G., Papakostas G.A., Kaburlasos, V.G.: A Novel Distance Measure of Intuitionistic Fuzzy Sets and Its Application to Pattern Recognition Problems, *International Journal of Intelligent Systems*, Vol. 27, 2012, pp. 396 – 409.

[12] Dubois D., Prade, H., A Review of Fuzzy Set Aggregation Connectives, *Inf. Sci.*, Vol. 36, 1985, pp. 85 – 121.

[13] Xu, Z., Intuitionistic Fuzzy Aggregation Operators, *IEEE Transactions on Fuzzy Systems*, Vol. 15, No. 6, 2007, pp. 1179 – 1187.

[14] Yager, R.R.: On Ordered Weighted Averaging Aggregation Operators in Multicriteria Decision Making, *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 18, No. 1, 1988, pp.183 – 190.

[15] Verma R., Sharma, B.D., On Generalized Intuitionistic Fuzzy Divergence (Relative Information) and Their Properties, *J. Uncertain Systems*, Vol. 6, No. 4, 2012, pp. 308 – 320.